# **Body-Centric Interactions With Very Large Wall Displays**

Garth Shoemaker<sup>1</sup>, Takayuki Tsukitani<sup>2</sup>, Yoshifumi Kitamura<sup>2</sup>, Kellogg S. Booth<sup>1</sup><sup>1</sup>Department of Computer Science<br/>University of British Columbia<sup>2</sup>Human Interface Engineering<br/>LaboratoryICICS/CS Building<br/>201-2366 Main Mall<br/>Vancouver, BC, Canada<br/>{garths,ksbooth}@cs.ubc.caSuita, Osaka 565-0871, Japan<br/>{tsukitani.takayuki,<br/>kitamura}@ist.osaka-u.ac.jp

## ABSTRACT

We explore a set of body-centric interaction techniques for very large wall displays. The techniques described include: virtual tools that are stored on a user's own body, protocols for sharing personal information between co-located collaborators, a shadow representation of users' bodies, and methods for positioning virtual light sources in the work environment. These techniques are important as a group because they serve to unify the virtual world and the physical world, breaking down the barriers between display space, personal body space, and shared room space. We describe an implementation of these techniques as integrated into a collaborative map viewing and editing application.

**ACM Classification:** H5.2 [Information interfaces and presentation]: User Interfaces. - Input devices and strategies.

General terms: Design, Human Factors

Keywords: Large displays, interaction techniques.

#### INTRODUCTION

For the last several decades, the dominant model of computing interaction has been a single user sitting at a desk, manipulating data shown on a small screen using a keyboard and a mouse. This model differs dramatically from how users work in the physical world. Users of physical data artifacts regularly use large tables and walls for information sharing and manipulation. These large surfaces are effective for tasks such as brainstorming [4], scheduling [26], and asynchronous data sharing [23]. In order to support these same tasks, it is desirable for computing systems to model the beneficial properties of large physical surfaces.

Ongoing advances in display hardware technology promise that within a relatively short time very large displays will be both practical and affordable. With the introduction of these devices a new model of interaction should be introduced to augment or replace the traditional modes of keyboard and



Figure 1: Users performing text input and sketching input with the system. The system was used to explore a variety of body-centric interaction techniques.

mouse. This new model of interaction may draw inspiration from familiar use of physical media including whiteboards and tabletops, but will also incorporate techniques only made possible by interactive computing systems. In this paper we explore a body-centric interaction approach that combines both familiar real-world elements and novel techniques.

Our work extends previous work by Shoemaker et al. [22], which focussed on the use of body shadows as a user embodiment to provide awareness and support long-distance reaching. Compared to this previous work we consider a more complete *model-based* description of the physical entities relevant to interaction: user, collaborators, displays, and room. A general model of these entities allows us to design interaction techniques that recognize and leverage the complex inter-relationships between all elements of the model.

Our first contribution is the implementation of a pipeline that builds models representing all of the physical entities relevant to interaction in a large wall display environment. This novel representation of users and displays located in space is general enough to be used as the basis for a variety of interaction techniques. Our second contribution is the implementation of three classes of interaction techniques built using the model. The first class comprises single-user, bodycentric techniques that utilize an individual's proprioceptive sense. Examples include body-based tools and body-based storage. The second class unifies a user's actions in body space with feedback provided in the display space. The feedback mechanisms for body-based storage are an example of this. The third class provides collaborative support based on real-world social cues and behaviors. Examples include mechanisms for exchanging information with collaborators, and protocols for enforcing privacy.

# **RELATED WORK**

An important trend in interaction design is the focus on reality-based interfaces. As described by Jacob et al. [12], *real world* based interaction techniques allow for the leveraging of important themes, including *naive physics, body awareness & skills, environment awareness & skills,* and *social awareness & skills.* These themes provide a basis for interaction shared by most people and most cultures. On the other hand, virtual interactions typically provide utility beyond the limitations of the physical interactions they mimic. Thus there is a tradeoff between replicating real world interactions with the goal of easing use, and providing expressive power and efficiency beyond what real world interactions support. Reality-based techniques hold particular relevance to large displays systems that model the affordances of physical surfaces.

# Large Display Interaction

Large Display interaction in general is a very active area of research. Researchers are exploring both tabletop and wall display systems. Specific areas of interest include understanding human use of large surfaces, display hardware advancements, and the design of novel interaction techniques.

Motivation for the development of large display systems is driven by the understanding that large physical surfaces play important roles in everyday life. For example, large vertical surfaces such as whiteboards and blackboards are ubiquitous in educational settings. It has been argued that the invention of blackboards in the early  $19^{th}$  century was one of the most important developments in educational technology [2]. More modern examples of large display use include activity coordination, such as in hospital trauma centers [26], personal information storage [23], and the support of brainstorming activities [4]. Considered as a whole, this work demonstrates that large work surfaces possess unique attributes that provide powerful support for certain classes of tasks.

A major hurdle in the evolution of large display systems has been developing the hardware to make such systems practical and affordable. Recently some major breakthroughs have been made in this area. Dietz and Leigh introduced DiamondTouch, a multi-user touch surface based on capacitive sensing [5]. More recently Han introduced multi-touch tracking using frustrated total internal reflection (FTIR) [9]. Building on this, researchers at Microsoft have explored rich interactions through the use of cameras for simultaneously tracking hands, objects, and users [11]. None of these technologies is clearly superior to the others, and it is likely that all of them will influence further developments.

A variety of interaction approaches have been developed specifically for large displays. These include gestural techniques, such as distant freehand pointing explored by Vogel and Balakrishnan [24]. Laser pointers have also been investigated for use with large wall displays [25], although it has been found that jitter inherent in laser pointers poses problems [18]. Direct touch is also widely used for input, providing a high level of awareness compared to mouse input [10], but limitations in human reach can lead to problems as displays become very large. Nevertheless, several compelling touch approaches have been presented, including TNT [15] and ShapeTouch [3].

# **Body-Based Interactions**

A user's body plays a central role in the reality-based interface paradigm, and several systems have employed users' bodies for the purpose of supporting interaction with large wall displays.

VIDEOPLACE, by Krueger et al. [14], introduced the concept of using the entire body for both input and feedback. A shadow of the user is shown on the large wall display, and the user's movements are reflected directly in the shadow. The shadow contour is employed to interact directly with onscreen objects. Shadows have continued to be a research topic of interest. The "Shadow Communication" system was developed to facilitate remote interactions [17], while "Shadow Reaching" supports distance interaction on very large wall displays [22]. Shadows as interaction metaphors hold promise when applied to large display systems not only because they provide an intuitive interaction proxy, but also because they provide powerful awareness cues, as has been shown by Pavani and Castiello [19].

What the body-centric literature has yet to investigate thoroughly is the variety of roles different body parts play in real world interaction. Existing literature, such as Shadow Reaching and VIDEOPLACE, focusses mostly on a body proxy to provide feedback. It would be beneficial to combine this with an exploration of how the physical body itself could be better tied into the input/feedback loop. The body has a number of very powerful properties, including the senses of proprioception and privacy, that can be utilized in the design of interaction techniques.

## DESIGN PRINCIPLES AND GOALS

A primary philosophy guiding the design of our interaction techniques is that of reality-based interactions, as introduced by Jacob et al. In particular, we have strived to emphasize *body awareness & skills* (BAS) in the design of our single user interactions, and *social awareness & skills* (SAS) in the design of our collaborative interactions.

A second goal in designing our interaction techniques was to preserve flexibility. It was necessary to integrate the techniques into a specific system for the purpose of demonstration, but we had as a goal that the techniques would be general enough that they could be applied to a wide variety of tasks that might be performed on a large wall display.

## **Body-Centric Interactions**

The interaction techniques we have developed are driven by geometric models of users in an environment. A user's model represents, as closely as possible, all geometric properties of the user: location in the room, limb poses, and relationships with other users and displays. This body-centric approach to



Figure 2: Display space, body space, and room space.

interaction is a natural extension to previous work [22] that focussed on a body shadow proxy as a feedback mechanism.

It is useful to note that other interaction techniques also implicitly rely on body models. Touch interaction, for example, models the fingers of a user as they touch an interactive surface. Gestural techniques rely on a model of the user's hands in mid-air. It can be argued that cursor movements indirectly model the motions of a user's hand as it moves the mouse. Common to these approaches is the lack of a model of the user's entire body, nor are the models situated in the environment. This limits techniques because they can only act on information available in the model. Our approach is to fully model the entire interactive environment, as closely as possible, in order to produce a rich basis upon which interaction techniques can be designed.

## **Unifying Interaction Spaces**

An underlying goal of our work, made possible by a complete model of users and the environment, is to unify various physical spaces in support of interaction. We have identified three relevant spaces: *display space*, *body space*, and *room space* (Figure 2). These differ in several measures, including geometric dimensionality, physicality, and the degree to which they are considered personal.

*Display Space:* As seen in Figure 2, display space has a dimensionality of two, at least for traditional flat non-stereo displays. Displays have a certain degree of physicality, but this physicality is restricted to the uniform flat surface of the display. In terms of its personal nature, large displays are frequently shared between multiple users, and thus should not be considered personal spaces. The use of small individual displays, which is beyond the scope of this paper, would introduce personal display spaces.

*Body Space*: Body space refers to the space inhabited by users' bodies. Its dimensionality can be seen in different ways. The volume inhabited by the body is three dimensional, the surface of the body is two dimensional, and the motions of the different limbs constrained by joints define various dimensions or degrees of freedom for motion. Body space is both highly physical and personal as a result of a user's senses (e.g. touch, proprioception). The personal element is emphasized by social rules of personal space [6].

It is important to note that the there are numerous reference frames associated with body space. As a user moves their limbs a proprioceptive awareness is maintained of the relative locations and orientations of these parts. The middle coordinate system shown in Figure 2 is specific to the user's left forearm, but there are other coordinate systems associated with each other limb. The significance of these different reference frames is that they constitute a rich system of relationships that can be easily managed and understood by the user and can be leveraged in the design of interaction techniques [13].

*Environment Space*: Environment space is defined as being the area of space adjacent to the display within which users locate themselves. It has a dimensionality of three. In terms of physicality, there is little interaction directly with the space itself, however, it can be considered to be physical due to the fact that it is inhabited by users' bodies. There is also a possibility for objects, either physical or virtual, to be located in the space. Environment space is not considered to be personal, as it is shared by all users.

We propose that display space and body space be unified through the use of a shadow embodiment on the display. This embodiment serves several purposes. First, it allows for the user to virtually inhabit the display space without being physically in contact with it. The user can be located anywhere in the environment space, and still be virtually situated in the display. Second, it allows for the user to make use of body space for the purpose of interacting with the display space. This is not normally possible, for example in the case of touch interaction, since interaction can only occur literally in the plane of the display. Thus, the shadow serves to unify display space with body space, and allows the user to take advantage of the rich interactive potential of personal body space for the purposes of interacting with the rich feedback potential of display space.

Display space also serves a purpose in unifying body space and environment space. Virtual objects in environment space, not being physical, cannot be interacted with literally. The display can be used, however, to provide feedback to the user who otherwise interacts directly with the virtual object. The user, modelled in the context of the environment, is free to move about the space to interact with virtual objects. If there are physical objects that are also modeled as part of the scene model, the user can also interact directly with those objects, and resultant changes can be reflected in the virtual world. Using this approach, environment space is essentially a direct extension of display space, an extrusion of display space into the third dimension. The main advantage of environment space is that the user is free to move about in it, while the main disadvantage is that feedback must be provided indirectly through display space.

By employing a complete scene model we are able to unify display space, body space, and environment space. With this unification we are able to design interaction techniques that best utilize the properties of the three spaces. We can combine the capabilities of display space, the proprioceptive and personal nature of body space, and the ability of environment space to act as 3D context to the virtual environment.



Figure 3: The sensing/modelling/rendering pipeline. From left to right: sensed locations of key joints, construction of an approximate skeleton based on known joint locations, and rendering of a human mesh and corresponding shadow. The images shown are from real data displayed in a test suite.

## SYSTEM IMPLEMENTATION

In order to develop a complete scene model the system follows the *sensing*, *modeling*, *rendering* approach described in [22]. Segmenting the architecture into three components introduces modularity into the design. A variety of sensing modules can be used, either independently or in conjunction, to produce a scene model and visualizations of the model through which body- and shadow-based interactions can be supported. The modules described in previous work, however, were limited, and only useful for supporting very basic shadow-based interactions.

We have developed more advanced modules for *sensing*, *modeling*, and *rendering* (Figure 3). These modules offer superior robustness, increased flexibility, and ultimately produce a much richer model, able to support interactions beyond what was previously possible.

#### Sensing

We have developed two different sensing modules. Both perform real-time measurement of points in space, typically associated with body joints. This position information is then fed into the modeling component, which builds a model of the users. In our implementation we track the two hands of each user, along with the shoulders.

The primary sensing module uses magnetic position sensors (Polhemus Liberty Latus) to track joint locations. Magnetic sensors do not suffer from occlusion problems, as no line-of-sight is required. Effective range from each tracking station is approximately 2 meters, however, a number of tracking stations can be placed over an area to increase coverage. The main disadvantage of magnetic markers is management of the active markers. They must be calibrated and batteries must be changed roughly every half hour. Because of these drawbacks this approach is suitable for the laboratory, but is not appropriate for deployment, at least with the current technology.

A second sensing module was developed, based on a vision approach. Coloured balls are attached to important joints of users, and a number of fixed cameras above the user track the balls. Triangulation is used to calculate the 3D locations of the balls. The main advantage of this approach is that the markers are passive, meaning there is no upkeep required. The cameras can be set to run continuously, and a user can start using the system without any calibration or initialization. There are two primary disadvantages to the approach. First, occlusion can be a major problem. Unless a large number of cameras are places at strategic locations, the markers will frequently be occluded. A second disadvantage is that tracking the markers is difficult when room lighting changes. Even when tracking in HSB space, unexpected changes in global brightness can cause major difficulties.

While both approaches are suitable for the laboratory, recent advances in consumer-grade sensing suggest that sensing of the kind necessary to develop accurate body models may soon be widely available. A primary example is Nintendo's Wii video game system, which utilizes accelerometers and infrared tracking to support gestural input. While this is not yet sufficient to produce an accurate model of the user, it is the first generation of such hardware, and it is likely that future generations of hardware will provide much more advanced sensing capabilities.

## Modeling

Our modeling component builds a virtual model of all relevant objects in the scene. This includes users and displays. The modeling module employs an inverse-kinematics (IK) approach and joint information provided by the sensing module. Known locations of a number of joints, along with length constraints for limbs and rotation limits for joints, are used with an inverse kinematic solver to derive a skeleton of the user. Without exhaustive instrumentation of the user's body it is not possible to produce an exact single solution to the IK equation. We have found that using only hand and shoulder positions provides enough information to produce a suitable approximation. For example, the calculated elbow location is usually close enough to not be noticeably different from actual position, and the lower body is not used extensively in our prototype, meaning a rough approximation of leg location is also sufficient.

Displays in the environment are assumed to be fixed, and thus we have not developed any means for real-time updating of these models. They are simply modeled as rectangles placed in the room. If the displays were mobile, approaches similar to those used to track and model users could be used to maintain a model of displays in the environment.

For generation of user shadows and shadow-based interac-



Figure 4: The mapping application used for investigating new interaction techniques.

tions, a model of the light sources in the room is also maintained. The system currently models one light source for each user, which follows a set of rules from one of several lighting behavior models. The lighting behavior models take as input all other models in the scene (i.e. displays, users, lights) and output the location of the light source associated with each user. Because light behavior models are associated with particular interactions, the specifics of the behaviors will be discussed later.

## Rendering

Rendering of virtual shadows is performed using the skeleton derived from the modeling component. The angles of deflection of the various joints of the skeleton are applied to a generic human body mesh to match the mesh to the skeleton. Once the human mesh has been posed correctly, rendering of the shadow can proceed using one of two implemented modules. The first module simulates a real-world shadow, but the second produces a Magic Lens-style shadow. Both rendering approaches are overlaid onto the initial rendering of the interactive content, whatever it may be.

A simulated real-world shadow is seen as a darkened region on top of the normally displayed content. This visualization is accomplished by rendering the human mesh to the pixel buffer as a series of alpha blended black polygons. The opacity of the triangles is kept constant in order to produce a consistent shadow colour. Back-face culling is also necessary, so that each shadow pixel is only darkened once, for a consistent shadow colour.

The Magic Lens style visualization is achieved using the stencil buffer. First, the stencil buffer is cleared to contain 0 for every pixel. Then the body mesh is rendered to the buffer as a series of 1s. This mask of 1 values can then be used to mask normal rendering operations to the pixel buffer. Any arbitrary content can be rendered, and it will be constrained to the the region defined by the shadow. A typical use of this approach is to render two different versions of the same content. For example, a map can be shown within the shadow, while satellite imagery is shown outside.

## APPLICATION CONTEXT: COLLABORATIVE MAPPING

When developing new interaction techniques it is desirable to investigate those techniques within the context of a realworld application. This ensures that any new techniques will be feasible to implement and practical to use.

We have developed a map viewing and editing application based on Google Maps (Figure 4). A map application is appropriate because a basic version will possess enough functionality to be useful to a casual user, but there are many possibilities for developing advanced functionality. The application supports a number of features that are useful for enabling exploration of interaction techniques:

*Panning + Zooming.* Users can pan the map and perform smooth zooming operations. We chose not to implement rotation because, unlike tabletop displays, orientation has special meaning on vertical surfaces: it is common for "up" to be mapped to north.

*Sketched Annotations.* The application allows for free-form sketched annotations. The annotations are geo-referenced, meaning all points are fixed on the map, and annotations grow as the user zooms in, and shrink as the user zooms out.

*Text Annotations.* There is also support for text entry using a soft keyboard. It has been shown that soft keyboards are effective for text entry on large displays, and can be used in conjunction with mid-air input approaches [21].

The application is capable of rendering several different 2D data layers on the large display, including satellite imagery and street maps.

#### SINGLE USER INTERACTION TECHNIQUES

As previously described, our approach of employing a scene model, including body models of all users, serves to unify display space, body space, and environment space. In this section we describe a number of interaction techniques that unify these different spaces in different ways.

## **Data Manipulation**

Manipulation of data shown on the display occurs through a perspective projection applied to the user's body. A virtual light source in environment space is associated with each user. As the user moves about, the virtual light source projects the user's body onto the display. General interaction with the data occurs through this projected shadow, triggered using a hand-held device. In our implementation we use Nintendo Wiimotes, one held in each hand, to trigger events. In this case, the virtual light source in environment space serves to unify body space and display space.

This approach to supporting data manipulation was previously described [22], and hence will not be explored in depth. It is worth noting, however, that data manipulation is now much richer, given the more advanced sensing and modeling described here, along with support for dynamic virtual light sources that will be expanded upon later.

#### **Tool Access**

Mode and tool selection on traditional computing platforms is typically performed either through the use of menus or but-



Figure 5: A user reaches his left hand towards his right shoulder to access a tool. This mechanism allows for immediate tool selection regardless of user location in the room, and leverages the proprioceptive sense.



Figure 6: A user's personal files are virtually stored in their physical chest. The user can navigate those files and pull files of interest into the shared workspace. Navigation is managed by referring to the large display.

tons shown on-screen. Fixed menus are not well-suited to very large displays. Because users move about the environment, a fixed menu will often not be within reach. Button pads are more suitable because they can be moved, but doing so requires explicit action. Instead we propose a body-centric method for tool and mode selection that overcomes the limitations of these other approaches.

The body-centric tool approach makes use of virtual tools stored at different locations on the user's body (Figure 5). The user accesses a tool or enables a mode by moving a hand to the appropriate location and clicking a button. This leverages body awareness & skills through its similarity to realworld actions such as picking a tool from a pocket or a belt. There are several advantages to this approach. First, no matter where the user is standing, all tools are readily available. Second, each user naturally possesses a personal set of tools; in collaborative settings there is no confusion regarding who "owns" what tools. Third, since tool location is registered relative to body joints (e.g. elbow, waist, knee), the user is able to utilize proprioception when accessing tools. Proprioception allows users to access known tools without relying on visual feedback on the display, or remembering arbitrary gestures. This can speed up access, and minimize the requirement for showing feedback on the display, which might distract other users.

In our implementation, to select a tool a user first presses the thumb button on the controller. This enters tool selection mode, and tool feedback is shown on the display. The user then moves the controller to the part of the body where the desired tool is located and presses the trigger button. The desired tool is then selected. It is not necessary for the user to refer to the visual feedback, but this information is useful for situations where the user has not yet learned the body location of the desired tool.

The advantages of proprioception have been recognized by other researchers. For example, Boeck et al. explored the "object in hand" metaphor, which combined force feedback with proprioception for object manipulation [1]. Mine et al. also utilized proprioception in the context of fully immersive environments [16].

## **Personal Data Access**

There are many situations in which a user would want to access personal information in a collaborative work setting. For example, a user may have personal notes written in a text file that need to be accessed, or may have relevant photos that should be shared. This technique provides mechanisms for accessing and sharing personal information.

Personal information storage and access uses a metaphor similar to body-based tools. Each user's torso serves as a virtual container of infinite size, from which personal data files can be accessed (Figure 6). This virtual storage is mapped to a user's personal computer or network drive. A user can use his or her hands to open, expand, and search through files virtually stored in the torso. This approach has many of the same benefits of body-based tools. First, personal files are always in close proximity and readily accessible to the owner, and second, there is little possibility for confusion regarding who "owns" which storage area.

There are several other advantages that are specific to the torso storage technique. First, centering the navigation on the torso also centers it between the user's arms. This makes it easy for the user to interact with the data, which is important because navigation through a complex file space is not a trivial task. Second, the torso is simultaneously the most massive part of a person's body, and the center of the person's body. The mass of the torso lends itself to being a metaphorical container for vast amounts of information. The fact that it is central to the whole body makes it a very personal part of the body, which also associates well with the personal nature of the data being accessed.

Visual feedback is provided through a data browsing widget in the form of a familiar hierarchical file browser shown in a grid layout. This is a suitable general purpose solution, however, if the application deals with only specific kinds of personal data a more special-purpose widget could be designed.



Figure 7: C/D gain depends on the relative distances of the light source and the user from the display. With the two positioned orthogonal to the display the C/D gain is defined as  $y = \frac{x}{x-1}$  where y = C/D Gain and x = light/user distance ratio.

#### **Dynamic Light-Source Positioning**

A single virtual light source in environment space is associated with every user, and the projection of the user from the light source location is used to support interaction. Allowing for dynamic positioning of the light is valuable, first because it can allow the user to reach arbitrary locations on the screen, and second because altering the location of the light can be used to adjust the control-display (C/D) input gain (Figure 7). This customizability is an advantage over either touch or laser-pointer input, where C/D gain is not configurable. Quite often there is a need to strike a balance between support for distance reaching and accurate manipulation by limiting the C/D gain. Because the light is virtual it can be assigned arbitrary behavior, either algorithmically defined, or controlled manually by a user. We implemented three different light-source behaviors (Figure 8), suitable for different situations.

*User Following.* This is an algorithmically defined behavior. Based on the known location of the user's shoulders, the system calculates the cross product of the vector between the shoulders and the "up" vector, producing a vector that points out the back of the user. The vector is scaled to a desired length, and the virtual light source is placed at the end of the vector. The result is that the light source follows the user around as if it were attached to a "tail". This behavior is useful because no matter how the user moves or turns, the light source is always appropriately positioned to project the user's shadow directly in front of the user. This can be advantageous, as the user can turn left or right in order to reach long distances over a wide display surface.

The length of the virtual "tail" behind the user can be determined by a number of rules. In the simplest case it is of fixed length. This can be limiting, however, because the shadow may become impractically large if the user moves too far away from the display. A superior solution is to relate the length of the tail to the distance of the user from the display.

*Orthographic.* This behavior is also algorithmically defined. It depends on the location of the user, and on the location and



Figure 8: A user's virtual light source can be positioned based on different behaviors. A composite of three such behaviors are shown colour coded. Green: *user following*, Red: *orthographic*, and Yellow: *manually positioned*. Note that the colours are for illustrative purposes only.

surface normal of the primary display. The system calculates a vector normal to the display surface, in the direction of and in line with the user. The light source is placed at a very large distance in the direction of that vector. The result is a near-orthographic projection of the user's shadow onto the primary display.

The purpose of this behavior is to provide a shadow mode of minimal distortion, with minimal opportunity for confusion. Distortion is minimized when the projection is orthogonal to the plane of the display. Confusion is minimized as a result of the projection being located directly in front of the user. Close proximity minimizes the chance that the shadow will interfere with operations being performed by other users.

*Manually Positioned.* At times users may wish to manually position a light source. This may occur if, for example, the user wishes to optimize shadow projection for interaction with a particular region of a very large display. A manually positioned light source also provides a very stable projection, which may make detailed work simpler.

A variety of approaches can be taken for supporting user control of the light source. In our application the user points in the direction where the shadow is to appear. The light source is then positioned behind the user in the direction opposite to the direction pointed. The distance  $d_l$  between the light source and the user is a function of the distance  $d_h$  of the user's hand to the user's body. Because the user is restricted by arm length, the distance can be exaggerated by the system. For example:  $d_l = d_h^2 + c$ .

Behavior Transitioning. This is a not a behavior on its own, but rather a means of transitioning between other behaviors. When switching from one behavior to another it is undesirable for the light source to jump instantly from one position to another. This can cause confusion for the user and collaborators. Instead, the system transitions from the position calculated by the old behavior function  $p = f_o$  to the position calculated by the new behavior  $p = f_n$  over a short period of time T by calculating a linear blend of the two functions



Figure 9: As user 1 enters user 2's outer and inner collaboration threshold light sources for the users transition from *user following* to *orthographic* behavior.

 $p = (1 - t/T) * f_o + (t/T)f_n$ . This provides continuity of the shadow projection.

# COLLABORATIVE INTERACTION TECHNIQUES

Large display systems are frequently used for co-located collaboration. It is desirable for such systems to seamlessly support natural collaborative interactions. Although our current sensing and modeling approach only models the *geometric* properties of users and environments, it is possible to model some collaborative intentions based solely on user geometry.

## Synchronized Shadow Projections

The shape and location of each user's shadow depends on the projection applied to each user's body model, which is a direct result of the location of the user's personal light source. When two users are operating independently in their own regions of the workspace their light sources can be positioned independently without conflict, but when users are working together then it is desirable to coordinate light source locations such that collaborative interactions are efficient.

We identified several properties of shadows that are desirable when users are working in close collaboration. First, the shadows should be *consistent*, meaning the shadows should reflect a believable real-world placement of physical light sources. For example, if user A is standing to the left of user B, then user A's shadow should also be to the left of user B's shadow. Second, the shadows should *minimize conflict*. If users are close to one another then it is possible for shadows to overlap. The more shadows overlap, the more likely it is that users will be confused. These issues are not serious when users are working independently, but when they are collaborating closely they are can be important.

Our approach for synchronizing shadow projections is to monitor user locations in the environment. When the distance between users falls below some outer threshold, they are deemed to be initiating collaboration. This threshold distance can be set as desired, but a good value is somewhere in the range of 45cm-120cm, identified by Hall [8] as being a distance within which people will approach only if they are familiar with one another.

Once the users are judged to be in collaboration range the



Figure 10: Private data is shared by the literal action of passing it to the other user's hand.

system transitions to a lighting model consistent with the requirements. The orthographic lighting model fills those requirements, and as users approach one another each of their lights transitions to the new model. Once inside the inner threshold a consistent orthographic shadow visualization is shown for both users.

#### Access Control and Conflict Management

Consistent with our design guidelines emphasizing realitybased techniques, we designed our access control protocols to center around *Social Awareness & Skills*. It has been shown that much can be ascertained regarding the privacy of information artifacts based on how users treat those artifacts [20]. It has also been shown that systems can be designed to model and explicitly enforce these protocols [7]. We have extended these ideas to wall displays by designing explicit access control and conflict management protocols based on implicit assumptions about user actions, rather than explicitly being set by users.

A particularly important issue is the separation of private information, that which is not meant to be seen by others, from public information, that which can be freely seen by all collaborators. This is an issue with our system because it allows movement of information from a user's personal file space to the public shared wall workspace.

We enforce privacy by requiring all access to private data to take place in the literal body frame of reference, whereas access to public data takes place in the display's frame of reference. For example, in order for a user to move private data from body storage to the display, the user must first directly access that storage. Once the file has been moved to the shared display, however, it can be accessed in the display's frame of reference. In another scenario, if user A wants to grant user B permanent access to a personal file, the user must physically and literally pass the file to the other user's hand (Figure 10). Their hands must come in close proximity in order for the file to be passed. This protocol of forcing private information access to occur in body space builds on a person's sense of their own personal space. People will only enter another's personal space if permission has been given. Failure to achieve this permission will likely result in a defensive action, or at least an awareness that personal space is

being violated. As a result, the protocol helps guide access to personal information through the theme of *social awareness* & *skills*.

## **INFORMAL USER FEEDBACK**

We gathered preliminary user feedback from six users. Each user was shown the different features of the system, and was then given an opportunity to explore the system. For collaborative features the experimenter acted as the collaborating user. Notes were taken about user behavior during the sessions, and feedback was gathered both during and following the session. Each session lasted approximately half an hour.

All users were able to understand the basic concepts underlying the system. After one or two tries users understood the body-centric metaphor for tool selection, and similarly were able to navigate personal file space. With respect to the bodycentric approach, one user commented "You can't mess up." The different lighting behaviors were also easily understood, as were the collaboration protocols. This suggests that basing the interactions on real-world metaphors was a good decision. Nevertheless there were several lessons learned that can guide improvements to the system.

First, several participants commented that performance and realism are important in supporting the power of the shadow metaphor for interaction. The system exhibited occasional "hiccups," where there was an observable delay before rendering refresh. Furthermore, there were occasional sensing glitches, which caused the shadow to take on inaccurate poses. These events occasionally broke the users' mental model concerning the reality of the shadow representation. It will be important to improve both performance and shadow accuracy in order to help maintain the illusion of the shadow.

A second comment was that it was sometimes difficult to remember the state of the two different hands. Each hand can be in a different mode, which is more complex than normal desktop systems where only a single cursor mode has to be remembered. It was suggested that the visualization could be improved to help the user understand which hand is doing what.

Another comment centered around the physical device that was used. The way the Wiimote is designed to be held suggests that it is a pointing device, similar to a laser pointer. Unfortunately this is inconsistent with the shadow approach, and caused at least one user to attempt to activate a tool by *pointing* at a body part, instead of by *placing* the device at the body part. It is worth considering other input device that do not present the affordances of a pointing device.

# CONCLUSIONS AND FUTURE WORK

In this paper we have described a body-centric approach to supporting interaction with very large wall displays. This approach is inspired by the *reality-based* philosophy of interaction technique design, which allowed us to leverage themes such as *body awareness & skills* and *social awareness & skills*. The goal of this approach is to foster techniques that are, among other things, easy to learn, easily interpretable, and expressive.

We began by describing design principles that helped guide

our work. This included the description of various interaction spaces, including *display space*, *body space*, and *environment space*. These different spaces serve to segment the various frames of reference relevant to interaction with, and display of, different information artifacts.

We then described an implementation of a *sensing*, *modeling*, and *rendering* architecture that enabled our interaction technique development. This implementation was a significant advancement beyond implementation of the architecture described in previous work, allowing for real-time calculation of a geometric scene model describing users and displays in the context of a shared interactive environment.

Based on the implemented architecture and our design principles we were able to develop a number of body-centric interaction techniques appropriate for use with various large wall displays. These include single user techniques for storing virtual tools directly on a user's own body, a technique for accessing personal information based on the metaphor of a user's torso as a container, and a number of behavioral models for controlling the motion of a user's personal light source. We also developed several collaboration techniques, including a technique for synchronizing users' shadows to ease collaborative work, and a number of protocols for enforcing access control and managing conflict.

An important next step in our work is to support the development of a more fine-grained body model. While our model is holistic, in the sense that it represents the user's entire body in the context of the environment, it is not a very detailed model. Of particular importance is a more accurate model of the user's hands and fingers. Many existing interaction techniques rely on manipulation using individual fingers. We would like to integrate these techniques with our wholebody techniques. This would unify previous hand-specific work with our whole-body work in a beneficial manner. This work will largely center around incorporating different sensing modules into our architecture.

Another important next step is to extend the model beyond the geometric properties of the scene. Models of mental process and intent could be very useful in guiding interaction techniques. We have made initial steps in this direction by modeling some collaborative protocols, but there is much work left. This work will center around developing new modeling modules.

Finally, we will continue developing new body-centric interaction techniques. This will involve the design of both new means of manipulation and corresponding feedback mechanisms. We may adapt our existing and future techniques with special consideration for multiple display environments, possibly including handheld devices and tabletop displays.

## ACKNOWLEDGEMENTS

Funding for this research was provided by the Natural Sciences and Engineering Research Council of Canada through operating and equipment grants, scholarships, and the NEC-TAR strategic research network, and by Defence Research and Development Canada.

# REFERENCES

- 1. Boeck, J. D., Cuppens, E., Weyer, T. D., Raymaekers, C., and Coninx, K. Multisensory interaction metaphors with haptics and proprioception in virtual environments. In *Proceedings of NordiCHI '04* (2004), pp. 189–197.
- Buxton, B. Surface and tangible computing, and the "small" matter of people and design. In *IEEE International Solid-State Circuits Conference Digest of Technical Papers* (2008), pp. 24–29.
- Cao, X., Wilson, A. D., Balakrishnan, R., Hinckley, K., and Hudson, S. E. Shapetouch: Leveraging contact shape on interactive surfaces. In *Proceedings of TABLETOP '08* (2008), pp. 139–146.
- Cherubini, M., Venolia, G., DeLine, R., and Ko, A. J. Let's go to the whiteboard: How and why software developers use drawings. In *Proceedings of CHI '07* (2007), pp. 557–566.
- Dietz, P., and Leigh, D. Diamondtouch: a multi-user touch technology. In *Proceedings of UIST '01* (2001), pp. 219–226.
- 6. Felipe, N. J., and Sommer, R. Invasions of personal space. *Social Problems 14*, 2 (1966), 206–214.
- 7. Greenberg, S., and Boyle, M. Moving between personal devices and public displays. In *Workshop on Handheld CSCW, CSCW '98* (1998).
- 8. Hall, E. T. *The Hidden Dimension*. Peter Smith Publisher Inc, 1992.
- Han, J. Y. Low-cost multi-touch sensing through frustrated total internal reflection. In *Proceedings of UIST* '05 (2005), pp. 115–118.
- Hornecker, E., Marshall, P., Dalton, N. S., and Rogers, Y. Collaboration and interference: awareness with mice or touch input. In *Proceedings of CSCW '08* (2008), pp. 167–176.
- Izadi, S., Hodges, S., Taylor, S., Rosenfeld, D., Villar, N., Butler, A., and Westhues, J. Going beyond the display: a surface technology with an electronically switchable diffuser. In *Proceedings of UIST '08* (2008), pp. 269–278.
- Jacob, R. J., Girouard, A., Hirshfield, L. M., Horn, M. S., Shaer, O., Solovey, E. T., and Zigelbaum, J. Reality-based interaction. In *Proceedings of CHI '08* (2008), pp. 201–210.
- Jeannerod, M., and Biguer, B. Référence égocentrique et espace représenté. *Rev Neurol (Paris)* 145, 8–9 (1989), 635–639.
- Krueger, M. W., Gionfriddo, T., and Hinrichsen, K. Videoplace – an artificial reality. In *Proceedings of CHI* '85 (1985), pp. 35–40.

- Liu, J., Pinelle, D., Sallam, S., Subramanian, S., and Gutwin, C. Tnt: improved rotation and translation on digital tables. In *Proceedings of Graphics Interface '06* (2006), pp. 25–32.
- Mine, M. R., Frederick P. Brooks, J., and Sequin, C. H. Moving objects in space: exploiting proprioception in virtual-environment interaction. In *Proceedings of SIG-GRAPH* '97 (1997), pp. 19–26.
- Miwa, Y., and Ishibiki, C. Shadow communication: system for embodied interaction with remote partners. In *Proceedings of CSCW '04* (2004), pp. 467–476.
- Myers, B. A., Bhatnagar, R., Nichols, J., Peck, C. H., Kong, D., Miller, R., and Long, A. C. Interacting at a distance: measuring the performance of laser pointers and other devices. In *Proceedings of CHI '02* (2002), pp. 33–40.
- Pavani, F., and Castiello, U. Binding personal and extrapersonal space through body shadows. *Nature Neuroscience* 7 (2003), 14–16.
- Scott, S. D., Carpendale, M. S. T., and Inkpen, K. M. Territoriality in collaborative tabletop workspaces. In *Proceedings of CSCW '04* (2004), pp. 294–303.
- Shoemaker, G., Findlater, L., Dawson, J. Q., and Booth, K. S. Mid-air text input techniques for very large wall displays. In *Proceedings of Graphics Interface '09 (to appear)* (2009).
- Shoemaker, G., Tang, A., and Booth, K. S. Shadow reaching: a new perspective on interaction for large displays. In *Proceedings of UIST '07* (2007), pp. 53–56.
- Tang, A., Lanir, J., Greenberg, S., and Fels, S. Supporting transitions in work: Informing large display application design by understanding whiteboard use. In *Proceedings of GROUP '09* (2009).
- Vogel, D., and Balakrishnan, R. Distant freehand pointing and clicking on very large, high resolution displays. In *Proceedings of UIST '05* (2005), pp. 33–42.
- Wienss, C., Nikitin, I., Goebbels, G., Troche, K., Göbel, M., Nikitina, L., and Müller, S. Sceptre: an infrared laser tracking system for virtual environments. In *Proceedings of VRST '06* (2006), pp. 45–50.
- Xiao, Y., Lasome, C., Moss, J., Mackenzie, C. F., and Faraj, S. Cognitive properties of a whiteboard: A case study in a trauma centre. In *Proceedings of ECSCW '01* (2001), pp. 259–278.